Favorite Sites of a Persistent Random Walk

Arka Ghosh* Steven Noren† Alexander Roitershtein‡

October 7, 2018

Abstract

We consider favorite (i.e., most visited) sites of the symmetric persistent random walk on $\mathbb{Z}$, a discrete-time process typified by the correlation of its directional history. We show that the cardinality of the set of favorite sites is eventually at most three. This is a generalization of a result by Tóth for a simple random walk, used to partially prove a longstanding conjecture by Erdős and Révész. The original conjecture asserting that for the simple random walk on integers the cardinality of the set of favorite sites is eventually at most two was recently disproved by Ding and Shen.

Keywords: favorite sites, most visited sites, local time, correlated random walks, discrete Ray-Knight theorems.

Mathematics Subject Classification (2010): 60G50, 60J10, 60J55.

1 Introduction

Let $\lambda \in \left[\frac{1}{2}, 1\right)$. Let $\{X_s\}_{s=1}^{\infty}$ be a discrete-time Markov chain on the state space $\{-1, 1\}$. X_1 is either 1 or -1 with equal probability, and for each $s > 1$, the Markov chain has the transition probabilities for values $c \in \{-1, 1\}$

$$\begin{aligned} P(X_s = c | X_{s-1} = c) &= \lambda, \\ P(X_s = -c | X_{s-1} = c) &= 1 - \lambda. \end{aligned} \tag{1}$$

Define the symmetric nearest-neighbor persistent random walk $\{S_t\}_{t=0}^{\infty}$ by

$$S_t := \sum_{s=1}^{t} X_s,$$

*Department of Statistics, Iowa State University, Ames, IA 50011, USA; e-mail: apghosh@iastate.edu

†School of Mathematics and Computer Science, Lake Superior State University, Sault Sainte Marie, MI 49783, USA; e-mail: snoren@lssu.edu

‡Department of Statistics, Texas A&M University, College Station, TX 77843, USA; e-mail: alexander@tamu.stat.edu

with the convention $S_0 = 0$ w.p.1. Intuitively, S_t is similar to a simple symmetric random walk on $\mathbb{Z}$, except the direction of the motion of S_t has a bias towards the same direction its previous step. As a matter of fact, if $\lambda = \frac{1}{2}$ is permitted, then the persistent random walk can be seen as a generalization of the simple random walk.

The earliest works on the persistent random walk come from [13, 29], who each introduced the model as a way of describing certain physical phenomena, and from [15, 18] in mathematical literature (see also [7, 3, 16, 22, 24] and references therein). The persistent random walk and other related processes have seen applications in other fields of physics, economics and biology, such as random collision models [26], ballistic diffusion [33], quantum random walks [20, 31], investment portfolio optimization [3], and the movement of animals [6], among others. Overviews of applications can be found in [8, 33]. A random environment version of the model has been discussed in [2, 26].

We begin this study on the persistent random walk with defining relevant notation. We define the *local time* of a site $x \in \mathbb{Z}$ at time $t > 0$ as the number of visits x receives from the walk up to time t:

$$L(x, t) := \#\{0 < s \leq t : S_s = x\}.$$

For every time t, we also define the set of *favorite sites*, that is, the sites of $\mathbb{Z}$ that have been visited by the random walk the most by time t:

$$\mathcal{K}(t) := \{y \in \mathbb{Z} : L(y, t) = \max_{z \in \mathbb{Z}} L(z, t)\}.$$

Since the range of S_t is finite at any given point in time, $\#\mathcal{K}(t) < \infty$ w.p.1 for any t. Note that there are only two ways in which $\mathcal{K}(t)$ could change from $\mathcal{K}(t-1)$: either the local time of a site outside of $\mathcal{K}(t-1)$ becomes a maximum local time at time t, in which $\#\mathcal{K}(t) = \#\mathcal{K}(t-1) + 1$, or one of the sites in $\mathcal{K}(t-1)$ receives one more visit at time t, in which $\#\mathcal{K}(t) = 1$.

Finally, we define the random variable $f(r)$ to be the number of times $\#\mathcal{K}(t)$ becomes $r \in \mathbb{N}$:

$$f(r) := \#\{t \geq 1 : S_t \in \mathcal{K}(t), \#\mathcal{K}(t) = r\}. \tag{2}$$

In the simple walk case ($\lambda = \frac{1}{2}$), it was shown that $f(1) = f(2) = \infty$ w.p.1 in [10] and [5]. In [10], [11] and [12], Erdős and Révész conjectured that $f(r)$ was finite w.p.1 for $r \geq 3$ (see [27] for a an overview). The conjecture was partially proven in [30], in which it was shown that $f(4)$ was finite w.p.1, hence $f(r)$ for $r \geq 5$ as well. Our main result in this chapter reveals that the set of favorite sites for the persistent random walk behaves similarly, regardless of the amount of local directional bias.

Theorem 1.1 (Main Theorem). *For any choice of* $\lambda \in (\frac{1}{2}, 1)$*,*

$$E(f(4)) < \infty.$$

In particular, $f(4) < \infty$ *w.p.1.*

This theorem extends the result found for the simple random walk in [30]. It's somewhat surprising of a result for the persistent case; for λ close to 1 the persistent walk will cover the same large intervals of integers with the same number of visits. One would presume that the intervals of favorite sites will stay large, but the theorem shows that over time, the number of favorite sites will eventually be bounded above by 3.

It was recently shown in [9] that for the simple random walk on integers, $f(3) = \infty$ w.p.1. Together with the result of Tóth [30] this establishes the phase transition in the behavior of $f(n)$ for the simple random walk. It is plausible that a similar phase transition happens for a general class of random motions on $\mathbb{Z}$. The present paper discusses an example of the process that turns out to be amenable to an adaptation of the approach of Tóth [30]. We believe that $f(3) = \infty$ w.p.1. for the persistent random walk as well, but leave this as a direction for future work (see the discussion in Section 7).

Our method of proof will follow closely to that of [30], as the framework of sojourn times provides naturally closed formulations in the extension into the directionally-dependent persistent processes. As such, much of the notation and the key lemmas will appear similar to as they did in [30], albeit under a new random process. However, the extension will not be trivial, as the simple walk case in [30] provided simplifications in the essential formulations that are absent in the persistent case. Our proof for the persistent random walk will utilize some deep results into the studies of probability theory, mathematical statistics, asymptotic analysis and the theory of hypergeometric functions. We hope that the work for this proof will pave the way for the study in the number of favorite sites for other processes outside of the realm of simple random walks.

2 Definitions

Before we begin to prove the theorem, we first need to establish the preliminary definitions and observations. First, we define the upcrossings and downcrossings, respectively, of a site x:

$$\begin{aligned} U(x,t) &:= \#\{0 < s \le t : S_s = x, S_{s-1} = x-1\}, \\ D(x,t) &:= \#\{0 < s \le t : S_s = x, S_{s-1} = x+1\}. \end{aligned}$$

A couple of things to note here: $U(x,t)$ and $D(x,t)$ can be seen as a partition of the total local time $L(x,t)$, in that $L(x,t) = U(x,t) + D(x,t)$. Also, $U(x,t)$ and $D(x,t)$ are related to each other given the relative position of S_t in the following way:

$$U(x,t) - D(x-1,t) = \mathbb{1}_{\{0<x\le S_t\}} - \mathbb{1}_{\{S_t<x\le 0\}}, \tag{3}$$

$$D(x,t) - U(x+1,t) = -\mathbb{1}_{\{0<x\le S_t\}} + \mathbb{1}_{\{S_t<x\le 0\}}. \tag{4}$$

Using (3) and (4), we can rewrite the local time all in terms of either upcrossings or downcrossings:

$$\begin{aligned} L(x,t) &= D(x,t) + D(x-1,t) + \mathbb{1}_{\{0<x\le S_t\}} - \mathbb{1}_{\{S_t<x\le 0\}} &(5) \\ &= U(x,t) + U(x+1,t) - \mathbb{1}_{\{0<x\le S_t\}} + \mathbb{1}_{\{S_t<x\le 0\}}. &(6) \end{aligned}$$

Next, we define the following stopping times for the upcrossings and downcrossings above: for any $x \in \mathbb{Z}$ and $k \ge 0$,

$$\begin{aligned} T^U_{x,k} &:= \inf\{t \ge 1 : U(x,t) = k\}, \\ T^D_{x,k} &:= \inf\{t \ge 1 : D(x,t) = k\}. \end{aligned}$$

We can use these stopping times to help partition $f(4)$ into infinite random variables based on the location and visiting direction of the new favorite sites in the following way:

$$\begin{aligned} u_x(4) &:= \sum_{t=1}^{\infty} \mathbb{1}_{\{\Delta_t=1, x\in\mathcal{K}_t, \#\mathcal{K}(t)=4\}} \\ &= \sum_{k=1}^{\infty} \mathbb{1}_{\{x\in\mathcal{K}(T^U_{x,k}), \#\mathcal{K}(T^U_{x,k})=4\}}. \end{aligned}$$

$$\begin{aligned} d_x(4) &:= \sum_{t=1}^{\infty} \mathbb{1}_{\{\Delta_t=-1, x\in\mathcal{K}_t, \#\mathcal{K}(t)=4\}}, \\ &= \sum_{k=1}^{\infty} \mathbb{1}_{\{x\in\mathcal{K}(T^D_{x,k}), \#\mathcal{K}(T^D_{x,k})=4\}}. \end{aligned}$$

From here, we can see that

$$f(4) = \sum_{x\in\mathbb{Z}} (u_x(4) + d_x(4))$$

Note that, due to the symmetry of our persistent random walk model (in the sense that for any $x \in \mathbb{Z}$ and $t \geq 0$, $P(S_t = x|S_0 = 0) = P(S_t = -x|S_0 = 0)$), $u_x(4)$ is equal in distribution to $d_{-x}(4)$ for any $x \in \mathbb{Z}$. Hence, for the expectation of $f(4)$, we only need to concern ourselves with the nonnegative sites:

$$E(f(4)) = 2\sum_{x=1}^{\infty} E(u_x(4)) + 2\sum_{x=0}^{\infty} E(d_x(4)). \tag{7}$$

We can prove $E(f(4))$ is finite by showing the series on the right-hand side of (7) are both finite. For the rest of this work, we will set out to prove the following:

$$\sum_{x=1}^{\infty} E(u_x(4)) = \sum_{x=1}^{\infty}\sum_{k=1}^{\infty} P\left(x \in \mathcal{K}(T^U_{x,k}), \#\mathcal{K}(T^U_{x,k}) = 4\right) < \infty. \tag{8}$$

The proof that $\sum_{x=0}^{\infty} E(d_x(4)) < \infty$ is a similar exercise left to the reader.

3 Ray-Knight Representation

Now we introduce the offspring distribution for a sequence of critical branching processes which will be vital in the theorem's proof. For every $t \geq -1$ and $i \geq 1$, consider the random variable $\zeta_{t,i}$ with distribution

$$P(\zeta_{t,i} = j) = \begin{cases} 1-\lambda & \text{if } j = 0 \\ \lambda^2(1-\lambda)^{j-1} & \text{if } j \geq 1. \end{cases} \tag{9}$$

A note about this distribution is that its expectation is 1 and its variance is $2\frac{1-\lambda}{\lambda}$, as the computation of a couple of geometric series reveals.

For any given positive site $x \geq 1$, this random variable will represent the number of times a persistent particle will move from $x+1$ to x until eventually returning to $x-1$. When the particle first moves rightward onto x, it has a $1-\lambda$ probability of going against its rightward bias and moving leftward to $x-1$. If the particle goes right instead, the particle will take an excursion before returning to x again, which includes a downcrossing from $x+1$. This time, the particle has a $1-\lambda$ probability of moving right and starting another excursion, or a λ probability of moving left and ending the "trials".

Whenever a particle visits x from the left again after arriving at $x-1$, the memory of the Markov chain that dictate the particle's transition probabilities does not include any of its previous excursions to the right of x. Thus, every trial of $(x+1)$-to-x downcrossings for each x-to-$(x-1)$ downcrossing will be independent and identically distributed with each other. So the number of downcrossings between two adjacent sites will be a Markov chain dependent on the number of downcrossings between the next lower pair of sites.

For each $t \geq 0$ and $i \geq 1$, we will make i.i.d. copies of $\zeta_{t,i}$, call them $\zeta^*_{t,i}$ and $\zeta'_{t,i}$. The motivation for these new random variables are slightly different from that of $\zeta_{t,i}$, but they will be used for similar representations. Fix $k \geq 0$ and $x \geq 1$.

First, we will define a Galton-Watson process Y_t with $\{\zeta_{t,i}\}_{t=-1,i=1}^{\infty}$ as the offspring it produces each generation. Define the initial state $Y_{-1} = k$ and, for each $-1 \leq t < \infty$, let $Y_{t+1} := \sum_{i=1}^{Y_t} \zeta_{t+1,i}$. Y_t is then a Markov chain with transition probabilities $\pi(i,j)$, given by

$$\begin{aligned} \pi(i,j) &:= P\left(Y_{t+1} = j | Y_t = i\right) \\ &= \begin{cases} \delta_{0,j}, & i = 0 \\ \left(\frac{\lambda^2}{1-\lambda}\right)^i (1-\lambda)^j \sum_{k=(i-j)_+}^{i-1} \binom{i}{k}\binom{j-1}{i-k-1}\left(\frac{1-\lambda}{\lambda}\right)^{2k}, & i \geq 1 \end{cases} \end{aligned} \tag{10}$$

where $(a)_+ := \max\{a, 0\}$. Note that the right-hand side of (10) is the calculated i-fold convolution of (9). As a note of interest, setting $\lambda = \frac{1}{2}$ in (10) will reduce the right-hand side to the equivalent transition probabilities seen in [23] and [19], due to the Chu-Vandermonde identity (see [25]).

Next, define a Galton-Watson process Z_t with $\{\zeta^*_{t,i}\}_{t=-1,i=1}^{\infty}$ offspring and one intruder particle entering each generation. Let $Z_0 = k$ be the initial state and, for each $0 \leq t \leq x-1$, let $Z_{t+1} := \sum_{i=1}^{Z_t+1} \zeta^*_{t+1,i}$. Then Z_t is also a Markov chain with transition probabilities $\rho(i,j)$ given by

$$\begin{aligned} \rho(i,j) &:= P(Z_{t+1} = j | Z_t = i) \\ &= \left(\frac{\lambda^2}{1-\lambda}\right)^{i+1} (1-\lambda)^j \sum_{k=(i+1-j)_+}^{i} \binom{i+1}{k}\binom{j-1}{i-k}\left(\frac{1-\lambda}{\lambda}\right)^{2k}. \end{aligned} \tag{11}$$

Before defining the final process in this set, we first need to define a new random variable η with distribution

$$P(\eta = j) = \begin{cases} \frac{1}{2} & \text{if } j = 0 \\ \frac{1}{2}\lambda(1-\lambda)^{j-1} & \text{if } j \geq 1 \end{cases}$$

This variable describes the number of downcrossings from 0 to -1 until a first visit to 1. This will be used to define the third Galton-Watson process Y'_t, with initial state $Y'_0 = Z_{x-1}$, $Y'_1 := \eta \cdot \delta_{\{S_1=-1\}} + \sum_{i=1}^{Y'_0} \zeta'_{1,i}$, and $Y'_{t+1} := \sum_{i=1}^{Y'_t} \zeta'_{t+1,i}$ for each $1 \leq t < \infty$. We exclude the calculation of the transition probabilities of Y'_t, as they are not needed for this proof.

With these three processes defined, we are now ready to build our Ray-Knight type representation of the local times of S_t. For each $y \in \mathbb{Z}$, define $\Delta_{x,k}(y)$ by

$$\Delta_{x,k}(y) := \begin{cases} Y_{y-x} & \text{if} \quad x-1 \leq y < \infty \\ Z_{x-y-1} & \text{if} \quad 0 \leq y \leq x-1 \\ Y'_{-y} & \text{if} \quad -\infty < y \leq 0 \end{cases} .$$

By this construction, we arrive at the Ray-Knight type representation for the downcrossings of the persistent walk:

$$(\Delta_{x,k}(y), y \in \mathbb{Z}) \stackrel{\mathcal{D}}{=} \left(D(T^U_{x,k+1}, y), y \in \mathbb{Z}\right), \tag{12}$$

in which $\stackrel{\mathcal{D}}{=}$ means equal in distribution. Plainly speaking, $\Delta_{x,k}(y)$ represents the random number of downcrossings into site y before the $(k+1)^{\text{th}}$ upcrossing to x for any $y \in \mathbb{Z}$.

Now define the following random variable for each $y \in \mathbb{Z}$:

$$\Lambda_{x,k}(y) := \Delta_{x,k}(y) + \Delta_{x,k}(y-1) + \mathbb{1}_{\{0<y\leq x\}}. \tag{13}$$

$\Lambda_{x,k}(y)$ serves as the local time of y stopped at $T^U_{x,k+1}$, based on (5). Hence, using (5) and (13), we get the following Ray-Knight representation:

$$(\Lambda_{x,k}(y), y \in \mathbb{Z}) \stackrel{\mathcal{D}}{=} \left(L(T^U_{x,k+1}, y), y \in \mathbb{Z}\right)$$

The following is a list of random variables and events that we will use for the more technical aspects of the main theorem's proof:

$$\widetilde{Y}_t := Y_t + Y_{t-1}, \qquad \widetilde{Z}_t := Z_t + Z_{t-1} + 1, \qquad \widetilde{Y}'_t := Y'_t + Y'_{t-1}.$$

$$\begin{aligned} \sigma_h &:= \inf\{t \geq 0 : Y_t \geq h\} \\ \omega &:= \inf\{t \geq 0 : Y_t = 0\} \\ \sigma'_h &:= \inf\{t \geq 0 : Y'_t \geq h\} \\ \omega' &:= \inf\{t \geq 0 : Y'_t = 0\} \\ \tau_h &:= \inf\{t \geq 0 : Z_t \geq h\} \end{aligned}$$

$$\begin{aligned} \widetilde{\sigma}_{h,0} &:= 0, & \widetilde{\sigma}_{h,i+1} &:= \inf\{t > \widetilde{\sigma}_{h,i} : \widetilde{Y}_t \geq h\}, \\ \widetilde{\sigma}_h &:= \widetilde{\sigma}_{h,1} & & \\ \widetilde{\sigma}'_{h,0} &:= 0, & \widetilde{\sigma}'_{h,i+1} &:= \inf\{t > \widetilde{\sigma}'_{h,i} : \widetilde{Y}'_t \geq h\}, \\ \widetilde{\sigma}'_h &:= \widetilde{\sigma}'_{h,1} & & \\ \widetilde{\tau}_{h,0} &:= 0, & \widetilde{\tau}_{h,i+1} &:= \inf\{t > \widetilde{\tau}_{h,i} : \widetilde{Z}_t \geq h\}, \\ \widetilde{\tau}_h &:= \widetilde{\tau}_{h,1} & & \end{aligned}$$

$$
\begin{aligned}
A_{h,p} &:= \left\{ \max_{1 \le t < \infty} \widetilde{Y}_t \le h, \#\{1 \le t < \infty : \widetilde{Y}_t = h\} = p \right\} \\
&:= \{\widetilde{\sigma}_{h,p} < \infty = \widetilde{\sigma}_{h,p+1},\ \widetilde{Y}_{\widetilde{\sigma}_{h,i}} = h \text{ for } i = 1, \cdots, p\} \\
A'_{h,p} &:= \left\{ \max_{1 \le t < \infty} \widetilde{Y}'_t \le h, \#\{1 \le t < \infty : \widetilde{Y}'_t = h\} = p \right\} \\
&:= \{\widetilde{\sigma}'_{h,p} < \infty = \widetilde{\sigma}'_{h,p+1},\ \widetilde{Y}'_{\widetilde{\sigma}'_{h,i}} = h \text{ for } i = 1, \cdots, p\} \\
B_{x,h,p} &:= \left\{ \max_{1 \le t < x} \widetilde{Z}_t \le h, \#\{1 \le t < x : \widetilde{Z}_t = h\} = p \right\} \\
&:= \{\widetilde{\tau}_{h,p} < x \le \widetilde{\tau}_{h,p+1},\ \widetilde{Z}_{\widetilde{\tau}_{h,i}} = h \text{ for } i = 1, \cdots, p\}
\end{aligned}
$$

Plainly speaking, $\widetilde{\sigma}_{h,i}$, $\widetilde{\sigma}'_{h,i}$ and $\widetilde{\tau}_{h,i}$ are the i^{th} hitting times of the interval $[h, \infty)$ by their respective processes, and ω and ω' are the extinction times of their respective processes. Furthermore, $A_{h,p}$, $A'_{h,p}$ and $B_{x,h,p}$ are the events that the respective processes hit exactly p times its maximum level h either before extinction or, in $B_{x,h,p}$'s case, before time x.

With these events defined and the Ray-Knight representation established, we arrive at the following expression for $E(u_x(4))$ for any x:

$$
\begin{aligned}
E(u_x(4)) = \sum_{p+q+r=3} \sum_{h=1}^{\infty} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} & P(A_{h,p} | Y_0 = h - k - 1) \\
& \times \pi(k, h - k - 1) \\
& \times P(B_{x,h,q}, Z_{x-1} = \ell | Z_0 = k) \\
& \times P(A'_{h,r} | Y'_0 = \ell),
\end{aligned}
$$

which then leads to an upper bound for the left-hand side of (8):

$$
\begin{aligned}
\sum_{x=1}^{\infty} E(u_x(4)) \le \sum_{p+q+r=3} \sum_{h=1}^{\infty} \sum_{k=0}^{\infty} & P(A_{h,p} | Y_0 = h - k - 1) \\
& \times \pi(k, h - k - 1) \\
& \times \left(\sum_{x=1}^{\infty} P(B_{x,h,q} | Z_0 = k) \right) \\
& \times \left(\sup_{\ell \ge 0} P(A'_{h,r} | Y'_0 = \ell) \right)
\end{aligned}
\tag{14}
$$

We now introduce bounds to the values in the right-hand side of (14). The first set of bounds comes in the form of a proposition, which shall be proven in the next section:

Proposition 3.1. *For any $\epsilon > 0$ there exists a finite constant $C < \infty$ such that for any $h \ge 1$ and $k \ge 0$:*

1. *For any* $p \geq 0$,

$$\sum_{x=1}^{\infty} P(B_{x,h,p}|Z_0 = k) \leq Ch \tag{15}$$

2. *If either* $k \in [(h - h^{1/2+\epsilon})/2, h + h^{1/2+\epsilon})/2]$ *or* $p \geq 1$ *holds, then*

$$P(A_{h,p}|Y_0 = k) \leq (Ch^{-1/2+\epsilon})^{p+1} \tag{16}$$

$$\sum_{x=1}^{\infty} P(B_{x,h,p}|Z_0 = k) \leq (Ch^{-1/2+\epsilon})^{p+1} h \tag{17}$$

Next is a lemma on the sum of some extreme values of $\pi(i,j)$. The proof of this lemma will be postponed until Section 6. For organizational purposes, we will begin an alphabetical ordering of the lemmas which will be proven in Section 6, starting with the following lemma.

Lemma A. *For any* $\epsilon > 0$*, there exist constants* $C, \gamma > 0$ *such that, for any* $h \geq 1$*,*

$$\sum_{k:|h-2k|>h^{1/2+\epsilon}} \pi(k, h-1-k) < C\exp(-\gamma h^{2\epsilon}).$$

Using Proposition 3.1 and Lemma A, we can bound the terms of the right-hand side of (14) for each fixed h. To show this, first fix h, p, q, r and $\epsilon \in (0, \frac{1}{10})$, then decompose the right-hand side of (14) into two series, one for values of k such that $|h - 2k| \leq h^{1/2+\epsilon}$ and the other for $|h - 2k| > h^{1/2+\epsilon}$. The bounds of each of these sums will be represented in the following lines as a left term of a sum and a right term, respectively.

For the case in which $r = 0$, we have through Proposition 3.1 and Lemma A

$$\begin{aligned}\sum_{x=1}^{\infty} E(u_x(4)) &\leq \sum_{h=1}^{\infty} (Ch^{-1/2+\epsilon})^{p+q+2} h + (Ch)(C\exp(-\gamma h^{2\epsilon})) \\ &\leq \sum_{h=1}^{\infty} C' h^{-3/2+5\epsilon} < \infty\end{aligned}$$

for a large enough $C' < \infty$. If $r > 0$, we have

$$\begin{aligned}\sum_{x=1}^{\infty} E(u_x(4)) &\leq \sum_{h=1}^{\infty} (Ch^{-1/2+\epsilon})^{p+q+r+3} h + (Ch)(C\exp(-\gamma h^{2\epsilon})) \\ &\leq \sum_{h=1}^{\infty} C' h^{-2+6\epsilon} < \infty\end{aligned}$$

for another large $C' < \infty$. This shows (8), which then completes the proof of Theorem 1.1.

4 Proof of Proposition 3.1

To prove Proposition 3.1, we rely primarily on four main lemmas, whose proofs will be reserved for Section 6 along with the proof of Lemma A. We shall continue the alphabetical labeling of the lemmas. For all of the lemmas, fix $\epsilon > 0$.

The first lemma shows that the jumps of the Markov chains Y_t and Z_t is unlikely to be greater than $h^{1/2+\epsilon}$ until the Markov chains reach h.

Lemma B. *Define the maximal jumps of Y_t and Z_t:*

$$\begin{aligned} M_h &:= \sup\{|Y_t - Y_{t-1}| : 1 \le t \le \sigma_h \wedge \omega\}, \\ N_h &:= \sup\{|Z_t - Z_{t-1}| : 1 \le t \le \tau_h\}. \end{aligned}$$

There exist two constants, $C < \infty$ and $\gamma > 0$, such that for any $h \ge 1$ and $k \ge 0$, we have

$$\begin{aligned} P(M_h > h^{1/2+\epsilon} | Y_0 = k) &< C\exp(-\gamma h^{2\epsilon}), \\ P(N_h > h^{1/2+\epsilon} | Z_0 = k) &< C\exp(-\gamma h^{2\epsilon}). \end{aligned}$$

The next lemma are bounds on the probabilities that $\widetilde{Y}_t$ and $\widetilde{Z}_t$ enter the interval $[h, \infty)$ at exactly h.

Lemma C. *There exists a constant $C < \infty$ such that for any $h \ge 1$ and $k \ge 0$, we have*

$$\begin{aligned} P(\widetilde{\sigma}_h < \infty, \widetilde{Y}_{\widetilde{\sigma}_h} = h | Y_0 = k) &< Ch^{-1/2+\epsilon} \\ P(\widetilde{Z}_{\widetilde{\tau}_h} = h | Z_0 = k) &< Ch^{-1/2+\epsilon} \end{aligned}$$

Next is a bound on the probability that $\widetilde{Y}_t$ does not enter $[h, \infty)$ before extinction, given that Y_0 is close to $h/2$.

Lemma D. *There exists a constant $C < \infty$ such that for any $h \ge 1$ and $k \in \left[\frac{h-h^{1/2+\epsilon}}{2}, \frac{h+h^{1/2+\epsilon}}{2}\right]$,*

$$P(\widetilde{\sigma}_h = \infty | Y_0 = k) < Ch^{-1/2+\epsilon}.$$

Finally, we give upper bounds to the expectation of the hitting times $\widetilde{\tau}_h$.

Lemma E. *There exists a constant $C < \infty$ such that for any $h \ge 1$ the following holds:*

1. *For any k,*
$$E(\widetilde{\tau}_h | Z_0 = k) < Ch.$$

2. *For $k \in \left[\frac{h-h^{1/2+\epsilon}}{2}, \frac{h+h^{1/2+\epsilon}}{2}\right]$,*
$$E(\widetilde{\tau}_h | Z_0 = k) < Ch^{1/2+\epsilon}.$$

Proof of Proposition 3.1. Using the strong Markov property of Y_t and Z_t, we arrive at the following recurrence relations for $p \ge 1$:

$$\begin{aligned} P(A_{h,p} | Y_0 = k) &= \sum_{\ell=0}^{\infty} P(\widetilde{\sigma}_h < \infty, Y_{\widetilde{\sigma}_h - 1} = h - \ell, Y_{\widetilde{\sigma}_h} = \ell | Y_0 = k) \times P(A_{h,p-1} | Y_0 = \ell), \\ \sum_{x=1}^{\infty} P(B_{x,h,p} | Z_0 = k) &= \sum_{\ell=0}^{\infty} P(Z_{\widetilde{\tau}_h - 1} = h - \ell, Z_{\widetilde{\tau}_h} = \ell | Z_0 = k) \times \left(\sum_{x=1}^{\infty} P(B_{x,h,p-1} | Z_0 = \ell) \right). \end{aligned} \tag{18}$$

Note that

$$\sum_{x=1}^{\infty} P(B_{x,h,0}|Z_0 = k) = \sum_{x=1}^{\infty} P(\widetilde{\tau}_h \geq x|Z_0 = k) = E(\widetilde{\tau}_h|Z_0 = k),$$

so we have the first part of the proposition when $p = 0$ by Lemma E.

Now consider the case $p = 1$. We divide the right-hand sides of (18) into two disjoint sums: one such that $\ell \in \left[\frac{h-h^{1/2+\epsilon}}{2}, \frac{h+h^{1/2+\epsilon}}{2}\right]$ and for all other values of ℓ. From Lemmas C and D, we have for the first sum of the first series

$$\sum_{\ell:|h-2\ell|\leq h^{1/2+\epsilon}}^{\infty} P(\widetilde{\sigma}_h < \infty, Y_{\widetilde{\sigma}_h-1} = h-\ell, Y_{\widetilde{\sigma}_h} = \ell|Y_0 = k)\times P(A_{h,0}|Y_0 = \ell) \leq \left(Ch^{-1/2+\epsilon}\right)\left(Ch^{-1/2+\epsilon}\right).$$

Also, from Lemma B, we have for the second sum

$$\begin{aligned}\sum_{\ell:|h-2\ell|>h^{1/2+\epsilon}}^{\infty} & P(\widetilde{\sigma}_h < \infty, Y_{\widetilde{\sigma}_h-1} = h-\ell, Y_{\widetilde{\sigma}_h} = \ell|Y_0 = k) \times P(A_{h,0}|Y_0 = \ell) \\ & \leq P(M_h > h^{1/2+\epsilon}|Y_0 = k) < C\exp(-\gamma h^{2\epsilon})\end{aligned}$$

As for the other relation, we obtain similar results using Lemmas B, C and E:

$$\sum_{\ell:|h-2\ell|\leq h^{1/2+\epsilon}}^{\infty} P(Z_{\widetilde{\tau}_h-1} = h-\ell, Z_{\widetilde{\tau}_h} = \ell|Z_0 = k)\times\left(\sum_{x=1}^{\infty} P(B_{x,h,0}|Z_0 = \ell)\right) \leq \left(Ch^{-1/2+\epsilon}\right)\left(Ch^{-1/2+\epsilon}\right),$$

$$\begin{aligned}\sum_{\ell:|h-2\ell|>h^{1/2+\epsilon}}^{\infty} & P(Z_{\widetilde{\tau}_h-1} = h-\ell, Z_{\widetilde{\tau}_h} = \ell|Z_0 = k) \times \left(\sum_{x=1}^{\infty} P(B_{x,h,0}|Z_0 = \ell)\right) \\ & \leq P(N_h > h^{1/2+\epsilon}|Z_0 = k)\left(\sup_{\ell\geq 0}\sum_{x=1}^{\infty} P(B_{x,h,0}|Z_0 = \ell)\right) \\ & < (C\exp(-\gamma h^{2\epsilon}))(Ch).\end{aligned}$$

These inequalities yield the second part of the proposition for $p = 1$. The cases of $p = 2, 3$ follow directly from the $p = 1$ case and from the recurrence relations in (18).

□

5 Preliminary Results on $\pi(i, j)$ and $\rho(i, j)$

Before we prove the lemmas introduced in Sections 3 and 4, we first need to establish some preliminary facts about the transition kernels $\pi(i, j)$ and $\rho(i, j)$ introduced in (10) and (11) respectively. For the simple random walk ($\lambda = \frac{1}{2}$), [19] and [23] showed that the variables Y_t and Z_t followed a negative binomial distribution, which was used to great effect in [30]. While the distributions of Y_t and Z_t in the persistent case are related to negative binomial distributions, there are enough differences to warrant a more meticulous kind of analysis.

The majority of the effort shown in this section will focus more on $\pi(i, j)$, as a proof of a result for $\pi(i, j)$ will closely resemble that for $\rho(i, j)$ with minor differences. However, analogous results of both kernels will be seen.

5.1 Expectation and variance

Observation 5.1. *For each $i \geq 0$ and $t \geq 1$,*

$$E(Y_t|Y_{t-1}=i)=i, \qquad \mathrm{Var}(Y_t|Y_{t-1}=i)=2\frac{1-\lambda}{\lambda}i \tag{19}$$

$$E(Z_t|Z_{t-1}=i)=i+1, \qquad \mathrm{Var}(Z_t|Z_{t-1}=i)=2\frac{1-\lambda}{\lambda}(i+1) \tag{20}$$

This is a trivial result, since both random variables are sums of i i.i.d. copies of the same variable $\zeta_{t,i}$ with distribution given in (9). Still, we will make use of these calculations in the next section.

5.2 Log-concavity

Here we explain the concept of logarithmic-concavity for sequences, which will be used to justify the unimodality of the distributions $\pi(i,\cdot)$ and $\rho(i,\cdot)$.

Definition 5.2. *A nonnegative sequence $\{a_k\}_{k=0}^{\infty}$ is* log-concave *if, for every $k \geq 1$,*

$$a_k^2 \geq a_{k-1}a_{k+1}.$$

If $\{a_k\}_{k=0}^{\infty}$ is a positive sequence, then this is equivalent to the sequence of ratios $\{\frac{a_{k+1}}{a_k}\}_{k=0}^{\infty}$ being nonincreasing.

For more information on the concept of log-concave sequences, we refer to [4], [28] and [32]. For now, we present this fact:

Theorem 5.3 (Corollary 3.3 in [32]). *The convolution of two log-concave sequences is also log-concave.*

Given the distribution in (9), it is a straightforward exercise to show that $\{\pi(1,j)\}_{j=0}^{\infty}$ and $\{\rho(1,j)\}_{j=0}^{\infty}$ are both log-concave sequences. Thus, since $\{\pi(i,j)\}$ is a convolution of $\{\pi(i-1,j)\}$ and $\{\pi(1,j)\}$ for each $i > 1$, $\{\pi(i,j)\}$ is log-concave for any $i \geq 1$ by mathematical induction. Similarly, $\{\rho(i,j)\}$ is log-concave as well. The log-concavity feature of these transition kernels ensures unimodality in the distribution, which is the next topic of discussion.

5.3 Mode of $\pi(i,j)$

Since $\{\pi(i,j)\}$ is unimodal, there exists $k \geq 0$ such that for all $j \geq 0$,

$$\pi(i,j) \leq \pi(i,k). \tag{21}$$

It is in our interests to find exactly which values this k could take for each fixed i and λ. While the exact values could depend on λ, we have found that they do not stray very far from the expectation of the random variable found in (19).

Theorem 5.4. *Fix $\lambda \in \left(\frac{1}{2}, 1\right)$ and $i \geq 1$. Suppose there is an integer k such that (21) holds. Then $k \in \{i-1, i\}$.*

Note that, by the log-concavity of $\{\pi(i,j)\}$ and the resulting monotonicity of $\left\{\frac{\pi(i,j+1)}{\pi(i,j)}\right\}$, we have the following corollary:

Corollary 5.5.

$$\pi(i,j-1) \leq \pi(i,j) \text{ if } j \leq i-1,$$
$$\pi(i,j+1) \leq \pi(i,j) \text{ if } j \geq i.$$

Our proof of Theorem 5.4 will require a reinterpretation of $\pi(i,j)$. First, allow us to define the Gauss hypergeometric function as in [1]:

Definition 5.6. *For each a, b and c, the* Gauss hypergeometric function $\,_2F_1(a,b;c;z)$ *is the function mapping $\{z : |z| < 1\}$ to $\mathbb{C}$ of the form*

$$\,_2F_1(a,b;c;z) := \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!},$$

where $(m)_n = \prod_{k=0}^{n-1}(m+k)$ is the Pochhammer symbol.

Define $z := \left(\frac{1-\lambda}{\lambda}\right)^2$, and observe that $z \in (0,1)$. Then the following representation can be formulated for $i \geq 1$ and $j \geq 0$:

$$\pi(i,j) = \begin{cases} \sqrt{z}^{j-i}(1-\sqrt{z})^{j+i} \binom{j-1}{i-1} \,_2F_1(j+1, j; 1+j-i; z), & i \leq j \\ \frac{\sqrt{z}^{i-j}}{(1+\sqrt{z})^{i+j}} \binom{i}{i-j} \,_2F_1(1-j, -j; 1+i-j; z), & i > j. \end{cases} \tag{22}$$

There are multiple ways of representing $\pi(i,j)$ with hypergeometric functions, particularly using the Euler transformation

$$\,_2F_1(a,b;c;z) = (1-z)^{c-a-b} \,_2F_1(c-a, c-b; c; z).$$

For the purposes of this proof, we shall use the representation in (22). In particular, we wish to observe the following instances of $\pi(i,j)$:

$$\pi(i,i+1) = i(1-\sqrt{z})^{2i}(\sqrt{z}-z) \,_2F_1(i+2, i+1; 2; z) \tag{23}$$

$$\pi(i,i) = (1-\sqrt{z})^{2i} \,_2F_1(i+1, i; 1; z) \tag{24}$$

$$\pi(i,i-1) = i\frac{\sqrt{z}}{(1+\sqrt{z})^{2i-1}} \,_2F_1(2-i, 1-i; 2; z) \tag{25}$$

$$\pi(i,i-2) = \frac{i(i-1)}{2} \frac{z(1+\sqrt{z})}{(1+\sqrt{z})^{2i-1}} \,_2F_1(3-i, 2-i; 3; z) \tag{26}$$

We recognize other transformations of the hypergeometric functions, in particular the following (incomplete) list of Gauss' contiguous relations. A note on notation: $F = \,_2F_1(a,b;c;z)$, $F(a\pm) = \,_2F_1(a \pm 1, b; c; z)$, and all other parameter changes use similar notation.

$$\begin{aligned} z\frac{ab}{c}F(a+,b+,c+) &= a(F(a+)-F) \\ &= b(F(b+)-F) \\ &= \frac{(c-b)F(b-)+(b-c+az)F}{1-z} \\ &= \frac{z}{c(1-z)}\left((c-a)(c-b)F(c+)+c(a+b-c)F\right) \end{aligned}$$

Using these contiguous relations, one can find the following equalities for every $i \geq 1$:

$$ {}_2F_1(i+2,i+1;2;z) = \frac{1-z}{z(2i+1)}\ {}_2F_1(i+2,i+1,1,z) - \frac{1}{z(2i+1)}\ {}_2F_1(i+1,i,1,z) \quad (27)$$

$$\begin{aligned} {}_2F_1(3-i,2-i;3;z) = \frac{2}{i(1+z)-2}\ {}_2F_1(2-i,1-i;2;z) \\ - \frac{(i-1)(1-z)+(2i-3)z^2}{(i-1)(i(1+z)-2)}\ {}_2F_1(3-i,2-i;2;z) \end{aligned} \quad (28)$$

Before moving on to the proof of Theorem 5.4, we need the following lemma:

Lemma 5.7.

1. *For all i, $c > 0$,*
$$\frac{{}_2F_1\left(i+2,i+1;c;z\right)}{{}_2F_1\left(i+1,i;c;z\right)} \leq \frac{1}{(1-\sqrt{z})^2}, \quad (29)$$

2. *For all $i \geq 3$ and $c > 0$,*
$$\frac{{}_2F_1\left(3-i,2-i;c;z\right)}{{}_2F_1\left(2-i,1-i;c;z\right)} \geq \frac{1}{(1+\sqrt{z})^2}. \quad (30)$$

Moreover, the left-hand side of each inequality converges to the right-hand side as $i \to \infty$.

Proof. The convergence to the right-hand side is a direct result of Theorem 2 of [14] on the asymptotic estimates of the large-parameter Gauss hypergeometric functions seen as solutions to given second-order recurrence relations, in particular on the $(++0)$ case. One can check that the inequalities hold for $i = 1$, and the inequalities for $i > 1$ comes from the monotonicity of the pointwise convergence. □

With (23)-(30), we can now prove Theorem 5.4.

Proof of Theorem 5.4. Since the log-concavity of $\pi(i,j)$ gives us that $\left\{\frac{\pi(i,j+1)}{\pi(i,j)}\right\}$ is nonincreasing, it is enough to show that $\frac{\pi(i,i-1)}{\pi(i,i-2)} > 1$ and $\frac{\pi(i,i+1)}{\pi(i,i)} < 1$. Using (23), (24), (27) and (29), we have the following:

$$\begin{aligned}
\frac{\pi(i,i+1)}{\pi(i,i)} &= i(\sqrt{z}-z)\frac{{}_2F_1\left(i+2,i+1;2;z\right)}{{}_2F_1\left(i+1,i;1;z\right)} \\
&= \frac{i}{2i+1}\frac{(1-z)(\sqrt{z}-z)}{z}\frac{{}_2F_1\left(i+2,i+1;1;z\right)}{{}_2F_1\left(i+1,i;1;z\right)} - \frac{i}{2i+1}\frac{\sqrt{z}-z}{z} \\
&\leq \frac{i}{2i+1}\frac{(1-z)(\sqrt{z}-z)}{z}\frac{1}{(1-\sqrt{z})^2} - \frac{i}{2i+1}\frac{\sqrt{z}-z}{z} \\
&= \frac{i}{2i+1}\left(\frac{(1-z)(\sqrt{z}-z)-(\sqrt{z}-z)(1-\sqrt{z})^2}{z(1-\sqrt{z})^2}\right) \\
&= \frac{i}{2i+1}\frac{2(\sqrt{z}-z)^2}{z(1-\sqrt{z})^2} \\
&= \frac{2i}{2i+1} < 1.
\end{aligned}$$

Also with (25), (26), (28) and (30), we get

$$\begin{aligned}
\frac{\pi(i,i-2)}{\pi(i,i-1)} &= \frac{i-1}{2}(\sqrt{z}+z)\frac{{}_2F_1\left(3-i,2-i;3;z\right)}{{}_2F_1\left(2-i,1-i;2;z\right)} \\
&= \frac{i-1}{i(1+z)-2}(\sqrt{z}+z) - \frac{(i-1)(1-z)+(2i-3)z^2}{2(i(1+z)-2)}(\sqrt{z}+z)\frac{{}_2F_1\left(3-i,2-i;2;z\right)}{{}_2F_1\left(2-i,1-i;2;z\right)} \\
&\leq \frac{i-1}{i(1+z)-2}(\sqrt{z}+z) - \frac{(i-1)(1-z)+(2i-3)z^2}{2(i(1+z)-2)}(\sqrt{z}+z)\frac{1}{(1+\sqrt{z})^2} \\
&= \frac{2(i-1)(\sqrt{z}+z)(1+\sqrt{z})-((i-1)(1-z)+(2i-3)z^2)\sqrt{z}}{2(i(1+z)-2)(1+\sqrt{z})} \\
&= 1-\frac{(2i-4)(1-z)+(i-3)\sqrt{z}(1-z)+(2i-3)z^{5/2}}{2(i(1+z)-2)(1+\sqrt{z})} \\
&< 1.
\end{aligned}$$

□

We get a similar result as Theorem 5.4 for $\rho(i,j)$, although we must account for the generational intruder particle of the Galton-Watson process Z_t.

Theorem 5.8. *Fix* $\lambda \in \left(\frac{1}{2},1\right)$ *and* $i \geq 1$. *Suppose there is an integer* k *such that* $\rho(i,j) \leq \rho(i,k)$ *for all* $j \geq 0$. *Then* $k \in \{i, i+1\}$.

5.4 Upper bound for $\frac{\pi(i,j-1)}{\pi(i,j)}$ for $i < j$

While Theorem 5.4 provides a lower bound for $\frac{\pi(i,j-1)}{\pi(i,j)}$ when $i < j$, we now seek an upper bound for the ratio. To achieve this, we continue our analysis of hypergeometric functions, but now in the context of an existing result in the statistical study of contingency tables. We define the *noncentral hypergeometric distribution* $\mathrm{Hyper}(M_1, M_2, N_1, N_2, \theta)$ with the following formula for $\max(0, M_1 - N_2) \leq x \leq \min(N_1, M_1)$:

$$P(X=x)=\frac{\binom{N_1}{x}\binom{N_2}{M_1-x}\theta^x}{\sum_{u=\max(0,M_1-N_2)}^{\min(N_1,M_1)}\binom{N_1}{u}\binom{N_2}{M_1-u}\theta^u},$$

where $X\sim \mathrm{Hyper}(M_1,M_2,N_1,N_2,\theta)$. While the noncentral hypergeometric distribution has no direct application in our model, we utilize an upper bound for its expectation given in line (5.2) in [21]:

$$E(X)\le \frac{-c+\sqrt{c^2+4\theta(1-\theta)i(i-1)}}{2(1-\theta)}, \tag{31}$$

where $c:=N_1+N_2-(N_1+M_1)(1-\theta)$. Using this inequality, we have the following lemma:

Lemma 5.9. *Let $i<j$. Let $X\sim \mathrm{Hyper}(M_1=i-1,M_2=j,N_1=i,N_2=j-1,\theta=\left(\frac{1-\lambda}{\lambda}\right))$. Then $E(X)<(1-\lambda)(j-1)$.*

Proof. By (31), $E(X)\le\frac{-c+\sqrt{c^2+4\theta(1-\theta)i(i-1)}}{2(1-\theta)}$, where $c=i+j-1-(2i-1)(1-\theta)$. Note that $c>(2i-1)\theta$, $i-1<2(1-\lambda)(i-1/2)+(2\lambda-1)(j-1)$ and $i\le j-1$ So

$$\begin{aligned}\theta i(i-1) &< \theta(j-1)((1-\lambda)(2i-1)+(2\lambda-1)(j-1))\\ &= (1-\lambda)(j-1)(2i-1)\theta+(1-\lambda)^2\frac{2\lambda-1}{\lambda^2}(j-1)^2\\ &< (1-\lambda)(j-1)c+(1-\theta)(1-\lambda)^2(j-1)^2\end{aligned}$$

Thus, $c^2+4\theta(1-\theta)i(i-1)<c^2+4(1-\theta)(1-\lambda)(j-1)c+4(1-\theta)^2(1-\lambda)^2(j-1)^2=(c+2(1-\theta)(1-\lambda)(j-1))^2$. Therefore, $E(X)\le\frac{-c+\sqrt{c^2+4\theta(1-\theta)i(i-1)}}{2(1-\theta)}<(1-\lambda)(j-1)$. □

We use this lemma to get an upper bound for $\frac{\pi(i,j-1)}{\pi(i,j)}$ when $i<j$.

Lemma 5.10. $\frac{\pi(i,j-1)}{\pi(i,j)}<\frac{1}{1-\lambda}\frac{j-i}{j-1}+1$ *for all $i<j$.*

Proof.

$$\begin{aligned}\frac{\pi(i,j-1)}{\pi(i,j)} &= \frac{\sum_{k=0}^{i-1}\binom{i}{k}(1-\lambda)^k\binom{j-2}{i-k-1}\lambda^{2(i-k)}(1-\lambda)^{j-1-(i-k)}}{\sum_{k=0}^{i-1}\binom{i}{k}(1-\lambda)^k\binom{j-1}{i-k-1}\lambda^{2(i-k)}(1-\lambda)^{j-(i-k)}}\\ &= \frac{1}{1-\lambda}\frac{\sum\binom{i}{h}\binom{j-1}{i-k-1}\left(\frac{j-i+k}{j-1}\right)\left(\frac{1-\lambda}{\lambda}\right)^{2k}}{\sum\binom{i}{h}\binom{j-1}{i-k-1}\left(\frac{1-\lambda}{\lambda}\right)^{2k}}\\ &= \frac{1}{1-\lambda}\frac{j-i}{j-1}+\frac{1}{1-\lambda}\frac{1}{j-1}\frac{\sum k\binom{i}{k}\binom{j-1}{i-k-1}\left(\frac{1-\lambda}{\lambda}\right)^{2k}}{\sum\binom{i}{k}\binom{j-1}{i-k-1}\left(\frac{1-\lambda}{\lambda}\right)^{2k}}\\ &= \frac{1}{1-\lambda}\frac{j-i}{j-1}+\frac{1}{1-\lambda}\frac{1}{j-1}E(X)\\ &< \frac{1}{1-\lambda}\frac{j-i}{j-1}+1,\end{aligned}$$

where $X\sim \mathrm{Hyper}(M_1=i-1,M_2=j,N_1=i,N_2=j-1,\theta=\left(\frac{1-\lambda}{\lambda}\right))$. The last inequality above comes from Lemma 14. □

6 Proof of Lemmas

We now go on to prove Lemmas A-E, which will complete the proof of Theorem 1.1 through Proposition 3.1.

6.1 Lemma A

Proof of Lemma A. Assume that $h \geq 2$ and $k \geq 1$, and let $n = h-2$ and $\ell = k-1$. Note that $\pi(k, h-1-k)$ can be interpreted as the probability that the walk leaves an arbitrary site k times to the left before visiting it a total of h times, given that the site's h^{th} visit from the walk came from the left side. For each $m \in \mathbb{N}$, define K_m to be the number of downcrossings from the site given m visits to the site. We can write $K_m = \sum_{t=1}^{m} J_t$ such that for each t, $J_t = 1$ if the t^{th} visit to the site is immediately followed by a downcrossing and $J_t = 0$ otherwise. For the persistent walk, it is clear that $\{J_t\}_{t=1}^{\infty}$ is a Markov chain on the state space $\{0, 1\}$ with $P(J_t = 0|J_{t-1} = 0) = P(J_t = 1|J_{t-1} = 1) = 1 - \lambda$ and $P(J_t = 0|J_{t-1} = 1) = P(J_t = 1|J_{t-1} = 0) = \lambda$. It can also be shown easily that the $\{J_t\}_{t=1}^{\infty}$ is stationary with uniform stationary distribution $\mu \equiv \frac{1}{2}$.

Using this new notation, we have the following:

$$\pi(k, h-1-k) = P\left(K_{n+1} = k|J_{n+1} = 1\right) = P\left(K_n = \ell\right).$$

Thus,

$$\sum_{k:|h-2k|>h^{1/2+\epsilon}} \pi(k, h-1-k) = P\left(|n - 2K_n| > (n+2)^{1/2+\epsilon}\right). \tag{32}$$

We now seek for an upper bound for the right-hand side of (32). To accomplish this, we use a functional central limit theorem for Markov chains from [17]. Let $f : \{0, 1\} \to \mathbb{R}$ be defined as $f(x) = x$. Then $E_\mu f := \int_{\{0,1\}} f(x)\mu(dx) = \frac{1}{2}$ and $E_\mu f^2 = \frac{1}{2} < \infty$. Also, since $\{J_t\}_{t=1}^{\infty}$ is a finite Markov chain, it is uniformly ergodic. Thus, by Theorem 9 in [17], we get the following weak convergence as $m \to \infty$:

$$\sqrt{m}\left(\frac{1}{m}\sum_{t=1}^{m} J_t - E_\mu f\right) = \frac{K_m - \frac{m}{2}}{\sqrt{m}} \Rightarrow N(0, \sigma_f^2),$$

where $N(0, \sigma_f^2)$ is a normal distribution with mean 0 and variance $\sigma_f^2 > 0$. Using this central limit theorem, we can show that, for any $\gamma < 1/2$, $E\left(\exp\{\gamma\left(2K_m - m\right)^2/m\}\right)$ converges as $m \to \infty$. Hence,

$$\sup_m E\left(\exp\{\gamma\left(2K_m - m\right)^2/m\}\right) < \infty.$$

Using (32) and Markov's inequality, we finally arrive at our result.

□

6.2 Overshooting Lemma

Before proving the remaining four lemmas, we first want to establish a rather important result in the study of favorite sites of the simple walk for the case of the persistent walk.

It was shown in [30] that the probabilities of the Markov chains Y_t and Z_t going past a threshold point by a given amount, conditioned on the processes reaching the threshold for the first time at this instant, is comparable to the probability that the Markov chain achieves the same amount of overshoot conditioned on the threshold being reached on the very first step. This allows one to find simple asymptotic bounds for the conditional moments of Y_t and Z_t, essential for the application of optional stopping theorems in the proofs ahead.

Here, we obtain the analogous result for the persistent case.

Lemma 6.1 (Overshooting Lemma)**.** *For any* $0 \leq k < h \leq u$ *the following overshoot bounds hold:*

$$P(Y_{\sigma_h} \geq u | Y_0 = k, \sigma_h < \infty) \leq P(Y_1 \geq u | Y_0 = h, Y_1 \geq h) = \frac{\sum_{v=u}^{\infty} \pi(h,v)}{\sum_{w=h}^{\infty} \pi(h,w)}$$

$$P(Z_{\tau_h} \geq u | Z_0 = k) \leq P(Z_1 \geq u | Z_0 = h, Z_1 \geq h) = \frac{\sum_{v=u}^{\infty} \rho(h,v)}{\sum_{w=h}^{\infty} \rho(h,w)}$$

Proof. For $1 \leq h \leq u$,

$$P(Y_{\sigma_h} \geq u | Y_0 = k, \sigma_h < \infty) = \sum_{l=0}^{h-1} P(Y_{\sigma_h - 1} = l | Y_0 = k, \sigma_h < \infty) \frac{\sum_{v=u}^{\infty} \pi(l,v)}{\sum_{w=h}^{\infty} \pi(l,w)},$$

$$P(Z_{\tau_h} \geq u | Z_0 = k) = \sum_{l=0}^{h-1} P(Z_{\tau_h - 1} = l | Z_0 = k) \frac{\sum_{v=u}^{\infty} \rho(l,v)}{\sum_{w=h}^{\infty} \rho(l,w)}.$$

Note that if the ratios of the right-hand side are bounded above by the case in which $l = h$, we'd get our desired inequalities, since the probabilities of the right-hand side partition their respective conditioned event. It is then enough to show that the ratios on the right-hand side are increasing in l.

Observe the following relations for $\pi(l,v)$:

$$\begin{aligned}
\pi(l,v)\pi(l+1,v+1) - \pi(l+1,v)\pi(l,v+1) &= \pi(l,v) \cdot (\pi(1,\cdot) * \pi(l,\cdot))(v+1) \\
&\quad - \pi(l,v+1) \cdot (\pi(1,\cdot) * \pi(l,\cdot))(v) \\
&= (1-\lambda)\pi(l,v)\pi(l,v+1) + \sum_{j=1}^{v+1} \lambda^2 (1-\lambda)^{j-1} \pi(l,v+1-j)\pi(l,v) \\
&\quad - (1-\lambda)\pi(l,v)\pi(l,v+1) - \sum_{j=1}^{v} \lambda^2 (1-\lambda)^{j-1} \pi(l,v-j)\pi(l,v+1) \\
&= \lambda^2 (1-\lambda)^v \pi(l,0)\pi(l,v) \\
&\quad + \sum_{j=1}^{v} \lambda^2 (1-\lambda)^{j-1} (\pi(l,v+1-j)\pi(l,v) - \pi(l,v-j)\pi(l,v+1))
\end{aligned}$$

The terms in the sum of the last line are nonnegative, by the log-concavity of $\pi(l,\cdot)$. Thus, $\frac{\pi(l+1,v)}{\pi(l,v)} \leq \frac{\pi(l+1,v+1)}{\pi(l,v+1)}$. Similarly, $\frac{\rho(l+1,v)}{\rho(l,v)} \leq \frac{\rho(l+1,v+1)}{\rho(l,v+1)}$.
So, for all $v < w$,

$$\pi(l,v)\pi(l+1,w) \geq \pi(l+1,v)\pi(l,w)$$

$$\rho(l,v)\rho(l+1,w) \geq \rho(l+1,v)\rho(l,w)$$

Hence, for all $0 \leq l < h \leq u$,

$$\sum_{v=h}^{\infty} \pi(l,v) \sum_{w=u}^{\infty} \pi(l+1,w) \geq \sum_{v=h}^{\infty} \pi(l+1,v) \sum_{w=u}^{\infty} \pi(l,w),$$

$$\sum_{v=h}^{\infty} \rho(l,v) \sum_{w=u}^{\infty} \rho(l+1,w) \geq \sum_{v=h}^{\infty} \rho(l+1,v) \sum_{w=u}^{\infty} \rho(l,w).$$

Thus, $\left\{ \frac{\sum_{w=u}^{\infty} \pi(l,w)}{\sum_{v=h}^{\infty} \pi(l,v)} \right\}_{l=0}^{h}$ is an increasing sequence, and so is $\left\{ \frac{\sum_{w=u}^{\infty} \rho(l,w)}{\sum_{v=h}^{\infty} \rho(l,v)} \right\}_{l=0}^{h}$. This completes the proof. □

Using the Overshooting Lemma, we obtain the following set of inequalities:

Corollary 6.2. *There exist constants C_1, C_2, C_3 and C_4 such that for any $0 \leq k < h$,*

$$E\left(Y_{\sigma_h} | Y_0 = k, \sigma_h < \infty\right) \leq \frac{\sum_{v=h}^{\infty} \pi(h,v)v}{\sum_{w=h}^{\infty} \pi(h,w)} \leq h + C_1 h^{1/2} \tag{33}$$

$$E\left(Y_{\sigma_h}^2 | Y_0 = k, \sigma_h < \infty\right) \leq \frac{\sum_{v=h}^{\infty} \pi(h,v)v^2}{\sum_{w=h}^{\infty} \pi(h,w)} \leq h^2 + C_2 h^{3/2} \tag{34}$$

$$E\left(Z_{\tau_h} | Z_0 = k\right) \leq \frac{\sum_{v=h}^{\infty} \rho(h,v)v}{\sum_{w=h}^{\infty} \rho(h,w)} \leq h + C_3 h^{1/2} \tag{35}$$

$$E\left(Z_{\tau_h}^2 | Z_0 = k\right) \leq \frac{\sum_{v=h}^{\infty} \rho(h,v)v^2}{\sum_{w=h}^{\infty} \rho(h,w)} \leq h^2 + C_4 h^{3/2} \tag{36}$$

6.3 Lemma B

To begin the proof of Lemma B, we start with an application of Corollary 6.2.

Sublemma 6.3. *There exists a constant $C < \infty$ such that for any $h \geq 1$ and $k \geq 0$,*

$$E(\sigma_h \wedge \omega | Y_0 = k) < Ch^2.$$

Proof. Let $\mathcal{F}_t$ be the sigma algebra generated by the set $\{Y_s\}_{s=0}^{t}$. Then

$$E\left(Y_{t+1}^2 - 2\frac{1-\lambda}{\lambda} \sum_{s=0}^{t} Y_s - \left[Y_t^2 - 2\frac{1-\lambda}{\lambda} \sum_{s=0}^{t-1} Y_s \right] \middle| \mathcal{F}_t \right) = Var(Y_{t+1}|\mathcal{F}_t) - 2\frac{1-\lambda}{\lambda} Y_t = 0,$$

since $Var(\zeta_{t+1,i}) = 2\frac{1-\lambda}{\lambda}$ for any i. So $Y_t^2 - 2\frac{1-\lambda}{\lambda} \sum_{s=0}^{t-1} Y_s$ is a martingale. Thus, by the Optional Stopping Theorem, using stopping time $\sigma_h \wedge \omega$, we have

$$k^2 = E\left(Y_{\sigma_h \wedge \omega}^2 - 2\frac{1-\lambda}{\lambda} \sum_{s=0}^{\sigma_h \wedge \omega - 1} Y \middle| Y_0 = k \right) \leq E(Y_{\sigma_h}^2 | Y_0 = k) - 2E(\sigma_h \wedge \omega | Y_0 = k),$$

since $Y_s \geq 1$ for $s < \omega$ w.p.1. Therefore,

$$2E(\sigma_h \wedge \omega | Y_0 = k) \leq E(Y_{\sigma_h}^2 | Y_0 = k, \sigma_h < \infty) P(\sigma_h < \infty | Y_0 = k) - k^2 < C_2 h^2,$$

by (34) of Corollary 6.2. This completes the proof. □

We state the following inequality on the tail probabilities of sums of i.i.d. random variables, which is proven in [34].

Sublemma 6.4 (Exponential Kolmogorov Inequality)**.** *Let ξ_j, $j \geq 1$, be i.i.d. random variables with $E\left(e^{\theta|\xi_j|}\right) < \infty$ for some $\theta > 0$ and $E(\xi_j) = 0$. Then for any $N > 0$ and $n \in \mathbb{N}$,*

$$P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j}\xi_j\right| > N\right) \leq e^{-\theta N}\left(E\left(e^{\theta\xi_j}\right)^n + E\left(e^{-\theta\xi_j}\right)^n\right)$$

Let $\xi_j = \zeta_j - 1$, where $P(\zeta_j = x) = \begin{cases} 1-\lambda & \text{if } x = 0 \\ \lambda^2(1-\lambda)^{x-1} & \text{if } x \geq 1 \end{cases}$. Then, for any t,

$$\begin{aligned}
E(e^{t\xi_1}) &= (1-\lambda)e^{-t} + \frac{\lambda^2}{1-(1-\lambda)e^t} \\
&= \frac{(1-\lambda)e^{-t} - 1 + 2\lambda}{1-(1-\lambda)e^t} \\
&= \frac{2e^{-t} - e^{-2t} - 2(1-\lambda) + (1-\lambda)e^{-t} + e^{-2t} - 2e^{-t} + 1}{1-(1-\lambda)e^t} \\
&= 2e^{-t} - e^{-2t} + \frac{e^{-2t} - 2e^{-t} + 1}{1-(1-\lambda)e^t}.
\end{aligned}$$

Note that for fixed t, the formula in the final line decreases with an increase in λ, so for $\lambda \in [\frac{1}{2}, 1)$, $E(e^{t\zeta_1})$ is maximized at $\lambda = \frac{1}{2}$, which is the simple walk case. So, by [30], assuming that $\lambda \in \left[\frac{1}{2}, 1\right)$, there is a constant $\theta_0 > 0$ such that for all $\theta \in [0, \theta_0)$, $E(e^{\theta\xi_1}) < e^{2\theta^2}$ and $E(e^{-\theta\xi_1}) < e^{2\theta^2}$. Using the Exponential Kolmogorov Inequality and choosing $\theta = N/(4n)$, we obtain the following:

Sublemma 6.5. *There is a constant such that for any $N > 0$ and $n \in \mathbb{N}$ satisfying $N/(4n) < \theta_0$,*

$$P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j}(\zeta_i - 1)\right| > N\right) \leq 2\exp(-N^2/(8n)).$$

Proof of Lemma B. We prove the first inequality here in detail. The proof of the second inequality is similar and is left for the reader. Choose $0 < \gamma < \frac{1}{16}$.

$$\begin{aligned}
P(M_h > h^{1/2+\epsilon}|Y_0 = k) \leq & P(M_h > h^{1/2+\epsilon}, \sigma_h \wedge \omega \leq h^2\exp(\gamma h^{2\epsilon})|Y_0 = k) \\
& + P(\sigma_h \wedge \omega > h^2\exp(\gamma h^{2\epsilon})|Y_0 = k)
\end{aligned}$$

For the first term on the right-hand side, we represent the Markov chain Y_t as the sum of i.i.d. random variables,

$$Y_{t+1} = \sum_{j=1}^{Y_t}\zeta_{t+1,j},$$

in order to obtain the following:

$$\begin{aligned}
&P(M_h > h^{1/2+\epsilon}, \sigma_h \wedge \omega \le h^2 \exp(\gamma h^{2\epsilon}) \\
&\le P\left(\max\left\{\max_{1\le j\le h}\left|\sum_{i=1}^{j}(\zeta_{t,i}-1)\right| : 1 \le t \le h^2\exp(\gamma h^{2\epsilon})\right\} > h^{1/2+\epsilon}\right) \\
&= 1 - \left(1 - P\left(\max_{1\le j\le h}\left|\sum_{i=1}^{j}(\zeta_{t,i}-1)\right| > h^{1/2+\epsilon}\right)\right)^{h^2\exp(\gamma h^{2\epsilon})} \\
&\le h^2\exp(\gamma h^{2\epsilon}) P\left(\max_{1\le j\le h}\left|\sum_{i=1}^{j}(\zeta_{t,i}-1)\right| > h^{1/2+\epsilon}\right).
\end{aligned}$$

Note that the last inequality above comes from the analytical fact that $1 - na \le (1-a)^n$ for $0 \le a \le 1$ and $n > 1$. From Sublemma 6.5, we get

$$P\left(\max_{1\le j\le h}\left|\sum_{i=1}^{j}(\zeta_{t,i}-1)\right| > h^{1/2+\epsilon}\right) \le 2\exp(-h^{2\epsilon}/8).$$

Hence, with $\gamma < \frac{1}{16}$, there is a constant $C > 0$ such that

$$P(M_h > h^{1/2+\epsilon}, \sigma_h \wedge \omega \le h^2\exp(\gamma h^{2\epsilon}) \le 2h^2\exp\left(\left(\gamma - \frac{1}{8}\right)h^{2\epsilon}\right) \le C\exp(-\gamma h^{2\epsilon}).$$

For the second term, we have from Sublemma 6.3 that there is a constant C such that

$$E(\sigma_h \wedge \omega)|Y_0 = k) \le Ch^2.$$

We get the following inequality after applying Markov's inequality:

$$P(\sigma_h \wedge \omega > h^2\exp(\gamma h^{2\epsilon})|Y_0 = k) \le C\exp(-\gamma h^{2\epsilon}).$$

Since both terms are bounded above by scalar multiples of $\exp(-\gamma h^{2\epsilon})$, the result follows. □

6.4 Lemma C

In order to prove Lemma C, we need the following sublemma:

Sublemma 6.6. *There exists a constant C s.t. for any $h \ge 1$ and $\ell \in \left[\frac{h-h^{1/2+\epsilon}}{2}, \frac{h+h^{1/2+\epsilon}}{2}\right]$,*

$$\frac{\pi(\ell, h-\ell)}{\sum_{m\ge h-\ell}\pi(\ell, m)} < Ch^{-1/2+\epsilon}$$

Proof. We make use of the inequalities found in Corollary 5.5. We shall split this proof into two cases. First, assume $\ell \in \left[\frac{h}{2}, \frac{h+h^{1/2+\epsilon}}{2}\right]$. Then $\ell > h - \ell$. Let $\{\zeta_k\}_{k=1}^{\ell}$ be a set of i.i.d. random variables with the same distribution as in (9). Recall that $E(\zeta_k) = 1$ and $\mathrm{Var}(\zeta_k) = 2\frac{1-\lambda}{\lambda}$. Let $\sigma^2 = \mathrm{Var}(X)$, and let Φ be the standard normal cdf. Then

$$\begin{aligned}
\pi(\ell, h-\ell) &\leq \pi(\ell, \ell-1) = P\left(\sum_{k=1}^{\ell} \zeta_k = \ell - 1\right) \\
&= \lim_{t\to 1^-} P\left(\sum_{k=1}^{\ell} \zeta_k \leq \ell - 1\right) - P\left(\sum_{k=1}^{\ell} \zeta_k \leq \ell - t\right) \\
&= \lim_{t\to 1^-} P\left(\frac{\sum_{k=1}^{\ell}(\zeta_k - 1)}{\sqrt{\ell}\sigma} \leq -\frac{1}{\sqrt{\ell}}\right) - P\left(\frac{\sum_{k=1}^{\ell}(\zeta_k - 1)}{\sqrt{\ell}\sigma} \leq -\frac{t}{\sqrt{\ell}}\right) \\
&\leq \lim_{t\to 1^-} \Phi\left(-\frac{1}{\sqrt{\ell}\sigma}\right) - \Phi\left(-\frac{t}{\sqrt{\ell}\sigma}\right) + \frac{C_1}{\sqrt{\ell}} \\
&= \frac{C_1}{\sqrt{\ell}}.
\end{aligned}$$

The last inequality above comes from the Berry-Esseen inequality. By the central limit theorem applied to $\{\zeta_k\}_{k=1}^{\infty}$, $\lim_{\ell\to\infty} \sum_{m\geq \ell} \pi(\ell, m) = \frac{1}{2}$. So there is a constant $C_2 > 0$ s.t.

$$\sum_{m\geq h-\ell} \pi(\ell, m) \geq \sum_{m\geq \ell} \pi(\ell, m) \geq C_2.$$

By these inequalities, we have our result for $\ell \in \left[\frac{h}{2}, \frac{h+h^{1/2+\epsilon}}{2}\right]$.

We now continue with the $\ell \in \left[\frac{h-h^{1/2+\epsilon}}{2}, \frac{h}{2}\right]$ case for the proof of the lemma. Note that $\ell < h - \ell$. Let $k = \lfloor h - \ell + h^{1/2-\epsilon} \rfloor$. Then we get the following inequalities, which will be described in more detail below:

$$\begin{aligned}
\frac{\pi(\ell, h-\ell)}{\sum_{m\geq h-\ell} \pi(\ell, m)} &\leq (k - h + \ell + 1)^{-1} \frac{\pi(\ell, h-\ell)}{\pi(\ell, k)} \\
&\leq (k - h + \ell + 1)^{-1} \left(\frac{\pi(\ell, k-1)}{\pi(\ell, k)}\right)^{k-h+\ell} \\
&< (k - h + \ell + 1)^{-1} \left(1 + \frac{1}{1-\lambda} \frac{k-\ell}{k-1}\right)^{k-h+\ell} \\
&\leq h^{-1/2+\epsilon} \left(1 + \frac{1}{1-\lambda} \frac{h^{1/2+\epsilon} + h^{1/2-\epsilon}}{h/2 + h^{1/2-\epsilon} - 1}\right)^{h^{1/2-\epsilon}} \\
&\leq h^{-1/2+\epsilon} \left(1 + \frac{2}{1-\lambda} h^{-1/2+\epsilon}\right)^{h^{1/2-\epsilon}} \\
&\leq \exp\left(\frac{2}{1-\lambda}\right) h^{-1/2+\epsilon}.
\end{aligned}$$

The first inequality comes from Corollary 5.5. The second comes from the log-concavity of $\{\pi(\ell, j)\}$. The third is Lemma 5.10. The fourth used $k \leq h - \ell + h^{1/2-\epsilon}$ and $\ell \geq$

$(h - h^{1/2+\epsilon})/2$. The fifth is due to the convergence of the base of the exponent above, as well as the monotonicity of that convergence. The final inequality relies on the exponential convergence of the power, as well as the monotonicity of that convergence.

□

Proof of Lemma C. We provide details of the proof of the first inequality and leave the similar details of the second inequality for the reader. First, observe that

$$P(\widetilde{\sigma}_h < \infty, \widetilde{Y}_{\widetilde{\sigma}_h} = h | Y_0 = k) \\ = \sum_{\ell=0}^{\infty} P(\widetilde{\sigma}_h < \infty, Y_{\widetilde{\sigma}_h - 1} = \ell, Y_{\widetilde{\sigma}_h} = h - \ell | Y_0 = k)$$

We split the infinite sum above into two sums, one for values of ℓ inside the interval $\left[\frac{h-h^{1/2+\epsilon}}{2}, \frac{h+h^{1/2+\epsilon}}{2}\right]$ and the other for values of ℓ outside the interval. For the first sum, we use Sublemma 6.6 to obtain

$$\sum_{\ell:|h-2\ell| \le h^{1/2+\epsilon}} P(\widetilde{\sigma}_h < \infty, Y_{\widetilde{\sigma}_h - 1} = \ell, Y_{\widetilde{\sigma}_h} = h - \ell | Y_0 = k) \\ = \sum_{\ell:|h-2\ell| \le h^{1/2+\epsilon}} P(\widetilde{\sigma}_h < \infty, Y_{\widetilde{\sigma}_h - 1} = \ell | Y_0 = k) \frac{\pi(\ell, h-\ell)}{\sum_{m=h-\ell}^{\infty} \pi(\ell, m)} \\ \le C h^{-1/2+\epsilon}.$$

For the second sum, we use Lemma B to obtain

$$\sum_{\ell:|h-2\ell| > h^{1/2+\epsilon}} P(\widetilde{\sigma}_h < \infty, Y_{\widetilde{\sigma}_h - 1} = \ell, Y_{\widetilde{\sigma}_h} = h - \ell | Y_0 = k) \\ P(M_h > h^{1/2+\epsilon} | Y_0 = k) < C \exp(-\gamma h^{2\epsilon}).$$

Therefore, we arrive at the result. □

6.5 Lemma D

For Lemma D, we require the following sublemma based on Corollary 6.2:

Sublemma 6.7. *There exists a constant $C < \infty$ such that for any $0 \le k < h$,*

$$P(\sigma_h = \infty | Y_0 = k) < \frac{h-k}{h} + C h^{-1/2}.$$

Proof. Let $\mathcal{F}_t$ be the sigma algebra generated by the set $\{Y_s\}_{s=0}^t$. Then

$$E(Y_{t+1} | \mathcal{F}_t) = E(Y_{t+1} | Y_t) = 1 \cdot Y_t,$$

since $E(\zeta_{t+1,i}) = 1$ for any i. So Y_t is a martingale. Thus, by the Optional Stopping Theorem, using stopping time $\sigma_h \wedge \omega$, we have

$$k = E(Y_{\sigma_h \wedge \omega} | Y_0 = k) = E(Y_{\sigma_h} | Y_0 = k, \sigma_h < \infty) P(\sigma_h < \infty | Y_0 = k) \le (h + C_1 h^{1/2}) P(\sigma_h < \infty | Y - 0 = k),$$

by (33) of Corollary 6.2. This completes the proof. □

Proof of Lemma D. Note that by Sublemma 6.7 and Lemma B, we get for two constants C_1, C_2,

$$\begin{aligned}P(\widetilde{\sigma}_h = \infty | Y_0 = k) \leq & P(\widetilde{\sigma}_h = \infty, M_h \leq h^{1/2+\epsilon} | Y_0 = k) + P(M_h > h^{1/2+\epsilon} | Y_0 = k) \\ & \leq P(\sigma_{(h+h^{1/2+\epsilon})/2} = \infty | Y_0 = k) + P(M_h > h^{1/2+\epsilon} | Y_0 = k) \\ & \leq C_1 h^{-1/2} + C_2 \exp(-\gamma h^{2\epsilon})\end{aligned}$$

for values of $k \in \left[\frac{h-h^{1/2+\epsilon}}{2}, \frac{h+h^{1/2+\epsilon}}{2}\right]$. □

6.6 Lemma E

In order to prove Lemma E, we first need upper bounds for the moments of τ_h.

Sublemma 6.8. *There exists a constant $C < \infty$ such that for any $0 \leq k < h$,*

$$E(\tau_h | Z_0 = k) < (h - k) + Ch^{1/2}.$$

Proof. Let $\mathcal{G}_t$ be the sigma algebra generated by the set $\{Z_s\}_{s=0}^t$. Then

$$E(Z_{t+1} - (t+1) | \mathcal{G}_t) = E(Z_{t+1} | Z_t) - (t+1) = Z_t - t,$$

since $E(\zeta^*_{t+1,i}) = 1$ for any i. So $Z_t - t$ is a martingale. Thus, by the Optional Stopping Theorem, using stopping time τ_h, we have

$$E(\tau_h | Z_0 = k) = E(Z_{\tau_h} | Z_0 = k) - k \leq h - k + C_3 h^{1/2},$$

by (35) of Corollary 6.2. This completes the proof. □

Sublemma 6.9. *There exists a constant $C < \infty$ such that for any $0 \leq k < h$,*

$$E(\tau_h^2 | Z_0 = k) < Ch^2.$$

Proof. Let $\mathcal{G}_t$ be the sigma algebra generated by the set $\{Z_s\}_{s=0}^t$. Then

$$E((t+1)^2 - 2(t+1)Z_{t+1} | \mathcal{G}_t) = (t+1)^2 - 2(t+1) - 2tZ_t - 2Z_t = t^2 - 2tZ_t + (1 - 2Z_t) \leq t^2 - 2tZ_t,$$

since $Z_t \geq 1$ for any t. So $t^2 - 2tZ_t$ is a supermartingale. Thus, by the Optional Stopping Theorem, using stopping time τ_h, as well as the Cauchy-Schwarz Inequality, we have

$$E(\tau_h^2 | Z_0 = k) \leq 2E(\tau_h Z_{\tau_h} | Z_0 = k) \leq 2\sqrt{E(\tau_h^2 | Z_0 = k)}\sqrt{E(Z_{\tau_h}^2 | Z_0 = k)}.$$

Therefore, by (36) of Corollary 6.2,

$$E(\tau_h^2 | Z_0 = k) \leq 4E(Z_{\tau_h}^2 | Z_0 = k) \leq 4C_4 h^2.$$

□

Proof of Lemma E.

$$E(\widetilde{\tau}_h | Z_0 = k) = E(\widetilde{\tau}_h \mathbb{I}_{(N_h \leq h^{1/2+\epsilon})} | Z_0 = k) + E(\widetilde{\tau}_h \mathbb{I}_{(N_h > h^{1/2+\epsilon})} | Z_0 = k).$$

Note that $\widetilde{\tau}_h \mathbb{I}_{(N_h \leq h^{1/2+\epsilon})} \leq \tau_{(h+h^{1/2+\epsilon})/2}$ and $\widetilde{\tau}_h \leq \tau_h$ w.p.1. Thus, with the Cauchy-Schwarz Inequality, we get

$$E(\widetilde{\tau}_h | Z_0 = k) \leq E(\tau_{(h+h^{1/2+\epsilon})/2} | Z_0 = k) + \sqrt{E(\tau_h^2 | Z_0 = k)}\sqrt{P(N_h > h^{1/2+\epsilon} | Z_0 = k)}.$$

Through Lemma B and Sublemmas 6.8 and 6.9, we arrive at the result. □

7 Future Work

In [5], it was shown that for the simple random walk, the typical favorite site is transient. This immediately implied that $f(1) = f(2) = \infty$ almost surely, for $f(r)$ defined in (2), and the fact was also used to prove that $f(3) = \infty$ almost surely as well in [9]. The transience of the favorite sites is a property we would also desire for the persistent random walk, particularly in examining the cases of $f(r)$ for $r = 2, 3$. We conjecture that the favorite sites of the persistent walk exhibit the same transience, based on our observations on the distribution of local times. However, this conjecture does not easily extend from the simple walk case in [5], as the result relies on the strong invariance principle of the local times between simple walks and Brownian motion. As no such strong invariance exists for the local times of directionally-reinforced random walks, a different approach is necessary for the case of the persistent walk.

References

[1] Abramowitz, M. and Stegun, I. A. (1964). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables (Vol. 55)*, Courier Corporation.

[2] Alili, S. (1999). Persistent random walks in stationary environment, *Journal of Statistical Physics, 94*(3-4), 469–494.

[3] Allaart, P. and Monticino, M. (2008). Optimal buy/sell rules for correlated random walks, *J. Appl. Probab. 45,*(1), 33–44.

[4] Bagnoli, M. and Bergstrom, T. (2005). Log-concave probability and its applications. *Economic theory*, *26*(2), 445–469.

[5] Bass, R. F. and Griffin, P. S. (1985). The most visited site of Brownian motion and random walk. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, *70*(4), 527–535.

[6] Bovet, P. and Benhamou, S. (1988). Spatial analysis of animals' movements using a correlated random walk model. *Journal of theoretical biology*, *131*(4), 419–433.

[7] Cénac, P., Le Ny, A., de Loynes, B., and Y. Offret (2018). Persistent random walks I: Recurrence versus transience. *Journal of Theoretical Probability, 31*(1), 232–243.

[8] Codling, E. A., Plank, M. J. & Benhamou, S. (2008). Random walk models in biology. *Journal of the Royal Society Interface*, *5*(25), 813–834.

[9] Ding, J. and Shen, J. (2016). Three favorite sites occurs infinitely often for one-dimensional simple random walk. (https://arxiv.org/abs/1612.01983) To appear in *Annals of Probability*.

[10] Erdős, P. and Révész, P. (1984). On the favourite points of random walks. In *Mathematical Structures-Computational Mathematics-Mathematical Modelling* (Sofia), *2* 152–157.

[11] Erdős, P. and Révész, P. (1987). Problems and results on random walks. In *Mathematical Statistics and Probability Theory*, 59–65. Springer Netherlands.

[12] Erdős, P. and Révész, P. (1991). Three problems on the random walk in $\mathbb{Z}^d$. *Studia Scientiarum Mathematicarum Hungarica*, *26*(2-3), 309–320.

[13] Fürth, R. (1917). Einige Untersuchungen über Brownsche Bewegung an einem Einzelteilchen. *Annalen der Physik*, *358*(11), 177–213.

[14] Gil, A., Segura, J., & Temme, N. (2007). Numerically satisfactory solutions of hypergeometric recursions. *Mathematics of computation*, *76*(259), 1449–1468.

[15] Goldstein, S. (1951) On diffusion by discontinuous movements, and on the telegraph equation. *Quart. J. Mech. Appl. Math., 4*, 129–156.

[16] Gruber, U. and Schweizer, M. (2006). A diffusion limit for generalized correlated random walks, *J. Appl. Prob. 43*, 60–73.

[17] Jones, G. L. (2004). On the Markov chain central limit theorem. *Probability Surveys*, *1*, 299–320.

[18] Kac, M. (1974) A stochastic model related to the telegraphers equation. *Rocky Mountain J. Math., 4*, 497–509. Reprinting of an article published in 1956, Papers arising from a Conference on Stochastic Differential Equations (Univ. Alberta, Edmonton, Alta., 1972).

[19] Knight, F. B. (1963). Random walks and a sojourn density process of Brownian motion. *Transactions of the American Mathematical Society*, *109*(1), 56–86.

[20] Konno. N. (2009). Limit theorems and absorption problems for one-dimensional correlated random walks. *Stochastic Models, 25*(1), 28–49.

[21] Kou, S. G. and Ying, Z. (1996). Asymptotics for a 2×2 table with fixed margins. *Statistica Sinica*, *6*, 809–829.

[22] Mauldin, R. D., Monticino, M., & von Weizscker, H. (1996). Directionally reinforced random walks. *Advances in Mathematics*, *117*(2), 239–252.

[23] Ray, D. (1963). Sojourn times of a diffusion process. *Illinois Journal of Mathematics*, *7*(4), 615–630.

[24] Renshaw, E. and and Henderson, R. (1981) The correlated random walk. *J. Appl. Probab., 18*(2):403–414.

[25] Roman, S. (2013). *The Umbral Calculus*, Courier Corporation.

[26] Szász, D. and Tóth, B. (1984). Persistent random walks in a one-dimensional random environment. *Journal of Statistical Physics*, *37*(1), 27–38.

[27] Shi, Z. and Tóth, B. (2000). Favourite sites of simple random walk. *Periodica Mathematica Hungarica*, *41*(1-2), 237–249.

[28] Stanley, R. P. (1989). Log-concave and unimodal sequences in algebra, combinatorics, and geometry. *Annals of the New York Academy of Sciences*, *576*(1), 500–535.

[29] Taylor, G. I. (1922). Diffusion by continuous movements. *Proceedings of the London Mathematical Society*, *2*(1), 196–212.

[30] Tóth, B. (2001). No more than three favorite sites for simple random walk. *Annals of Probability*, *29*(1), 484–503.

[31] Venegas-Andraca, S. E. (2012). Quantum walks: a comprehensive review. *Quantum Information Processing 11,*(5), 1015–1106.

[32] Wang, Y. and Yeh, Y. N. (2007). Log-concavity and LC-positivity. *Journal of Combinatorial Theory, Series A*, *114*(2), 195–210.

[33] Weiss, G. H. (2002). Some applications of persistent random walks and the telegrapher's equation. *Physica A: Statistical Mechanics and its Applications*, *311*(3), 381–410.

[34] Williams, D. (1991). *Probability with martingales*, Cambridge: Cambridge University Press.